Cancer Cell, Volume *28*

Supplemental Information

Thermodynamic Interrogation of the Assembly of a Viral Genome Packaging Motor Complex

Teng-Chieh Yang,[1] David Ortiz,[1] Lyn'Al Nosaka,[2] Gabriel C. Lander,[2] and Carlos Enrique Catalano[1,*]

[1]Department of Medicinal Chemistry, School of Pharmacy, University of Washington, Seattle, Washington; and [2]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, California

Supplemental Information

Thermodynamic Model for a Monomer - Oligomer Self-Association Equilibrium. An equilibrium between a monomer (*P*, the terminase protomer in the present study) and a higher order complex composed of "*n*" monomers ("n_{mer}" species, P_n) can be described as,

$$n * P \underset{}{\overset{K_A^*}{\rightleftharpoons}} P_n$$

where *P* and P_n are the equilibrium concentrations of the free monomer and the n_{mer} oligomer, respectively, and K_A^* is the macroscopic equilibrium association constant,

$$K_A^* = \frac{[P_n]}{[P]^n}$$

S1a

Upon rearranging this becomes,

$$[P_n] = K_A^* \bullet [P]^n$$

S1b

According to this simple model, the total species concentration in terms of the monomer, [P_T], is,

$$[P_T] = [P] + n \bullet [P_n]$$

S1c

and by substitution

$$[P_T] = [P] + n \bullet K_A^* \bullet [P]^n$$

S1d

The fraction of monomer free in solution $\left(f_P\right)$ and the fraction of monomer assembled into the oligomer $\left(f_{Pn}\right)$ at equilibrium are,

$$f_P = \left(\frac{[P]}{[P_T]}\right) = \left(\frac{[P]}{[P] + n[P_n]}\right) \qquad \textbf{S1e}$$

$$f_{Pn} = \left(\frac{n[P_n]}{[P_T]}\right) = \left(\frac{n[P_n]}{[P] + n[P_n]}\right) \qquad \textbf{S1f}$$

Substitution of [P_n] according to **Equation S1b** affords expressions for the fraction of monomer free in solution and that assembled into the oligomer as a function of [P] and K_A^*,

$$f_P = \frac{[P]}{[P] + n \cdot K_A^* \cdot [P]^n} \qquad \textbf{S1g}$$

$$f_{Pn} = \frac{n \cdot K_A^* \cdot [P]^n}{[P] + n \cdot K_A^* \cdot [P]^n} \qquad \textbf{S1h}$$

Linkage Model for Ligand Binding and Oligomer Assembly. **Figure S1** describes a simple model for a protein self-association equilibrium (monomer - n_{mer}) that is

thermodynamically linked to the binding of a ligand, *L*, to the protein. The nature of the ligand is general and can be another protein, polynucleotide, a small molecule, or salt. In this model, "*n*" copies of a protein monomer (*P*) self-assemble into an n_{mer} oligomer P_n. In the absence of the ligand this is described by the equilibrium constant K_1. The monomer can also bind "*m*" copies of a ligand to yield PL_m, described by the equilibrium constant K_3, which can then self-assemble to the n_{mer} oligomer to afford the $(PL_m)_n$ complex, described by the equilibrium constant K_4. To complete the thermodynamic cycle, the n_{mer} oligomer can bind *m* molecules of *L* to afford the $(P{\bullet}L_m)_n$ oligomer, described by the equilibrium constant K_2. In this simple model, the stoichiometry of *L* associated with each monomer is not changed upon assembly of the n_{mer} oligomer.

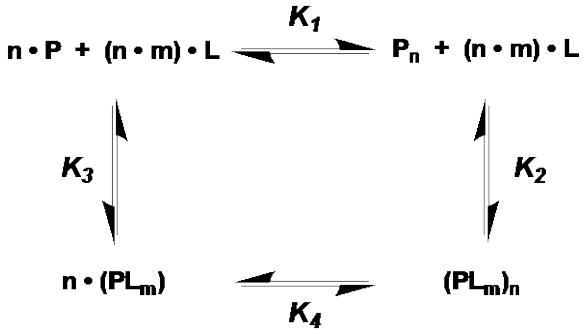

Figure S1. Protomer-Tetramer-Salt Linkage Model

Linkage Model for Salt Binding and Terminase Ring Tetramer Assembly. Our published data and the data presented in **Table 1** and **Figure 5** (*Main Text*) show that the macroscopic association constant for tetramer assembly, K_A^*, significantly increases with salt concentration. Here we derive an expression that describes the thermodynamic

linkage between salt binding and tetramer assembly (n= 4) to resolve the equilibrium binding constants based on the model presented in **Figure S1**. In this simple linkage model, each protomer binds m salt ions (L), both in isolation and in the context of the ring tetramer[1]. To link the observed salt effects to the experimentally determined macroscopic K_A^*, we first consider the situation where the concentration of protomer is such that $f_P = f_{P4}$ and according to **Equations S1g** and **S1h**,

$$\frac{[P]_{0.5}}{[P]_{0.5} + 4 \cdot K_A^* \cdot [P]_{0.5}^4} = \frac{4 \cdot K_A^* \cdot [P]_{0.5}^4}{[P]_{0.5} + 4 \cdot K_A^* \cdot [P]_{0.5}^4} \qquad \textbf{S3a}$$

which upon rearrangement yields,

$$[P]_{0.5} = \left(\frac{1}{4 \cdot K_A^*} \right)^{1/3} \qquad \textbf{S3b}$$

where $[P]_{0.5}$ is the *free* monomer concentration at $f_P = f_{P4}$. Substitution of this expression into **Equation S1d**,

$$[P_T]_{0.5} = \left(\frac{1}{4 \cdot K_A^*} \right)^{1/3} + 4 \cdot K_A^* \cdot \left(\left(\frac{1}{4 \cdot K_A^*} \right)^{1/3} \right)^4 \qquad \textbf{S3c}$$

which upon rearrangement yields,

[1] The model does not distinguish whether Na^+, Cl^-, or both ions affect protomer self-assembly.

$$[P_T]_{0.5} = 2 \cdot \left(\frac{1}{4 \cdot K_A^*} \right)^{1/3} \qquad \textbf{S3d}$$

where $[P_T]_{0.5}$ is the *total* monomer concentration under the condition where $f_P = f_{P4}$ described in terms of the experimentally determined macroscopic equilibrium constant K_A^*.

We next consider salt binding to the terminase protomer and to the ring tetramer within the context of the model presented in **Figure S1**, where n=4. By mass action, the total protomer concentration is,

$$[P]_T = [P] + [PL_m] + 4 \bullet [P_4] + 4 \bullet [(PL_m)_4] \qquad \textbf{S4a}$$

and according to the equilibria described in **Figure S1** it can be shown that,

$$[PL_m] = K_3^{1/4} \bullet [P] \bullet [L]^m \qquad \textbf{S4b}$$

$$[P_4] = K_1 \bullet [P]^4 \qquad \textbf{S4c}$$

$$[(PL_m)_4] = K_1 \bullet K_2 \bullet [P]^4 \bullet [L]^{4 \cdot m} \qquad \textbf{S4d}$$

and by substitution,

$$[P_T] = [P] + K_3^{1/4} \bullet [P] \bullet [L]^m + 4 \bullet K_1 [P]^4 + 4 \bullet K_1 \bullet K_2 \bullet [P]^4 \bullet [L]^{4 \cdot m} \qquad \textbf{S4e}$$

Next we take advantage of the fact that the fraction of protomer present as a tetramer species at equilibrium can be described by,

$$Y = \frac{4 \bullet [P_4] + 4 \bullet [(PL_m)_4]}{[P_T]} \qquad \textbf{S5a}$$

Under conditions where $f_P = f_{P4}$, Y= 0.5 and by substitution of expressions in **Equation S4b, S4c**, and **S4d** into **Equation S5a** one obtains,

$$0.5 = \frac{4 \bullet K_1 \bullet [P]^4 + 4 \bullet K_1 \bullet K_2 \bullet [P]^4 \bullet [L]^{4 \bullet m}}{[P] + K_3^{1/4} \bullet [P] \bullet [L]^m + 4 \bullet K_1 \bullet [P]^4 + 4 \bullet K_1 \bullet K_2 \bullet [P]^4 \bullet [L]^{4 \bullet m}} \qquad \textbf{S5b}$$

which upon rearrangement affords,

$$[P] = \left[P\right]_{0.5} = \left(\frac{1 + K_3^{1/4} \bullet [L]^m}{4 \bullet K_1 + 4 \bullet K_1 \bullet K_2 \bullet [L]^{4 \bullet m}} \right)^{1/3} \qquad \textbf{S5c}$$

Substitution of **Equation S5c** into **Equation S4e**, one obtains an expression for $[P_T]_{0.5}$ as a function of K_1, K_2, K_3, m, and the concentration of L:

$$[P_T]_{0.5} = \left(\frac{1 + K_3^{1/4} \bullet [L]^m}{4 \bullet K_1 + 4 \bullet K_1 \bullet K_2 \bullet [L]^{4 \bullet m}} \right)^{1/3} + K_3^{1/4} \bullet \left(\frac{1 + K_3^{1/4} \bullet [L]^m}{4 \bullet K_1 + 4 \bullet K_1 \bullet K_2 \bullet [L]^{4 \bullet m}} \right)^{1/3} \bullet [L]^m$$

$$+ 4 \bullet K_1 \bullet \left(\left(\frac{1 + K_3^{1/4} \bullet [L]^m}{4 \bullet K_1 + 4 \bullet K_1 \bullet K_2 \bullet [L]^{4 \bullet m}} \right)^{1/3} \right)^4 + 4 \bullet K_1 \bullet K_2 \bullet \left(\left(\frac{1 + K_3^{1/4} \bullet [L]^m}{4 \bullet K_1 + 4 \bullet K_1 \bullet K_2 \bullet [L]^{4 \bullet m}} \right)^{1/3} \right)^4 \bullet [L]^{4 \bullet m} \qquad \textbf{S5d}$$

Finally, by equating **Equation S3d**, which describes $[P_T]_{0.5}$ in terms of the experimentally measured macroscopic equilibrium constant K_A^*, with **Equation S5d**, which describes $[P_T]_{0.5}$ in terms of the linked equilibrium constants described in **Figure S1**, one obtains **Equation S5e** (presented as ***Equation 4 in the Main Text***):

$$2 \bullet \left(\frac{1}{4 \bullet K_A^*}\right)^{1/3} = \left(\frac{1 + K_3^{1/4} \bullet [L]^m}{4 \bullet K_1 + 4 \bullet K_1 \bullet K_2 \bullet [L]^{4 \bullet m}}\right)^{1/3} + K_3^{1/4} \bullet [L]^m \bullet \left(\frac{1 + K_3^{1/4} \bullet [L]^m}{4 \bullet K_1 + 4 \bullet K_1 \bullet K_2 \bullet [L]^{4 \bullet m}}\right)^{1/3}$$

$$+ 4 \bullet K_1 \bullet \left(\frac{1 + K_3^{1/4} \bullet [L]^m}{4 \bullet K_1 + 4 \bullet K_1 \bullet K_2 \bullet [L]^{4 \bullet m}}\right)^{4/3} + 4 \bullet K_1 \bullet K_2 \bullet [L]^{4 \bullet m} \bullet \left(\frac{1 + K_3^{1/4} \bullet [L]^m}{4 \bullet K_1 + 4 \bullet K_1 \bullet K_2 \bullet [L]^{4 \bullet m}}\right)^{4/3}$$

S5e

For the NLLS analysis, all of the parameters (K_1, K_2, K_3 and m) are allowed to float at any given concentration of L, in our case – the NaCl concentration, until the NLLS analysis finds the best fit of K^*.

We note that equation S5e is general and can be used to analyze any monomer-n_{mer} equilibrium reaction,

$$2 \bullet \left(\frac{1}{n \bullet K_A^*}\right)^{1/(n-1)} = \left(\frac{1 + K_3^{1/n} \bullet [L]^m}{n \bullet K_1 + n \bullet K_1 \bullet K_2 \bullet [L]^{n \bullet m}}\right)^{1/(n-1)} + K_3^{1/n} \bullet [L]^m \bullet \left(\frac{1 + K_3^{1/n} \bullet [L]^m}{n \bullet K_1 + n \bullet K_1 \bullet K_2 \bullet [L]^{n \bullet m}}\right)^{1/(n-1)}$$

$$+ n \bullet K_1 \bullet \left(\frac{1 + K_3^{1/n} \bullet [L]^m}{n \bullet K_1 + n \bullet K_1 \bullet K_2 \bullet [L]^{n \bullet m}}\right)^{n/(n-1)} + n \bullet K_1 \bullet K_2 \bullet [L]^{n \bullet m} \bullet \left(\frac{1 + K_3^{1/n} \bullet [L]^m}{n \bullet K_1 + n \bullet K_1 \bullet K_2 \bullet [L]^{n \bullet m}}\right)^{n/(n-1)}$$

S5f

Table S1. All data collected in Buffer containing 100 mM NaCl.

	Protomer $s_{(20,w)}$	**Tetramer** $s_{(20,w)}$	**Protomer-Tetramer** K_A^* (M^{-3})	$K_{A,app}$ §	$K_{D,app}$ †
Isolated Species ¶	5.12 ± 0.03	14.0 ± 0.2	-	-	-
$<s_{(20,w)}>$ Fit	-	-	7.50 (6.20, 9.20) x 10^{15}	1.95 x 10^5 M^{-1}	5.1 µM
SEDPHAT	5.28 (5.23, 5.32)	14.77 (14.58, 14.96)	-	-	-
SedAnal	5.23 (5.23, 5.25)	15.66 (15.58, 15.71)	7.63 (7.34, 7.93) x 10^{15}	1.97 x 10^5 M^{-1}	5.1 µM
Sed Eq Data	-	-	1.72 (1.59, 1.85) x 10^{16}	2.58 x 10^5 M^{-1}	3.9 µM

§ $K_{A,app} = \sqrt[3]{K_A^*}$

† $K_{D,app} = \left(1 / K_{A,app} \right)$

¶ Published Data: Maluf, Yang, and Catalano (2005) *J Mol Biol* 347:523-542; Maluf, et al. (2006) *Biochemistry* 45:15259-15268.

Table S2. All experiments contained terminase at a concentration equivalent to 1 μM protomer.

Salt	Fraction Tetramer
NaCl (500 mM)	0.78
NaOAc (500 mM)	0.70
KCl (500 mM)	0.73
K Glutamate (500 mM)	0.69
$MgCl_2$ (167 mM)	0.70
$MgSO_4$ (125 mM)	0.69

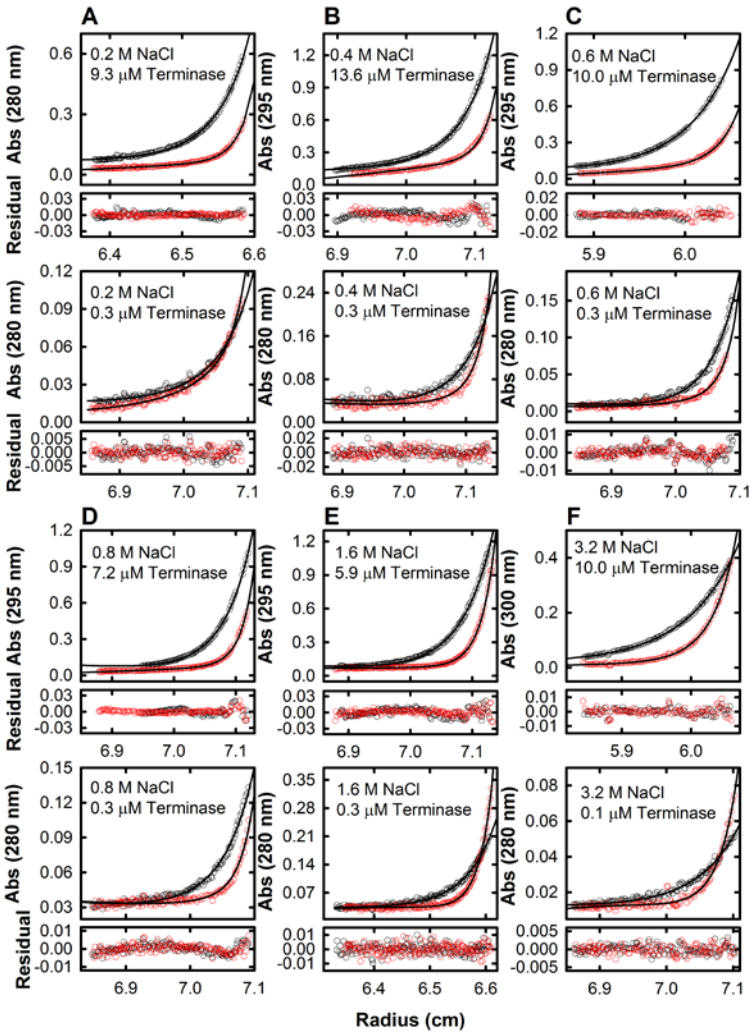

Figure S2. NaCl dependence of terminase self-association. Absorbance traces of terminase sedimentation equilibrium at different NaCl conditions: *Panels A-F* represents experiments carried out at 0.2, 0.4, 0.6, 0.8, 1.6 and 3.2 M NaCl, respectively. At each NaCl concentration, two concentrations of terminase were sedimented at two different speeds (8,000 and 12,000 RPM, black and red, respectively) and sedimentation was monitored at the indicated wavelength. The ensemble of data were analyzed according to a reversible monomer-tetramer equilibrium model as described in Materials and Methods holding *n* constant at 4, which affords a good fit (solid black lines with residuals shown below the equilibrium data).